

CLAIMS

- 1 1. A method for processing source data from a plurality
2 of diverse sources in a selected data domain, comprising:
3 specifying a unified schema that lists markup tags
4 in the selected data domain that can exist in a document
5 in the markup language;
6 defining correspondences of data fields from the
7 sources to the markup tags listed by the schema; and
8 mapping the source data in accordance with the
9 correspondences to generate unified data in the markup
10 language.
- 1 2. A method according to claim 1, wherein the markup
2 language comprises Extensible Markup Language (XML).
- 1 3. A method according to claim 2, wherein specifying
2 the unified schema comprises specifying a Document Type
3 Definition (DTD).
- 1 4. A method according to claim 2, wherein defining the
2 correspondences comprises defining data transformation
3 rules in Extensible Style Language (XSL).
- 1 5. A method according to claim 4, wherein mapping the
2 source data comprises transforming the data using an XSL
3 engine.
- 1 6. A method according to claim 1, wherein defining the
2 correspondences comprises selecting one or more of the
3 data fields in the sources to correspond to one of the
4 markup tags in the schema, and determining a conversion
5 function to apply to the one or more data fields.
- 1 7. A method according to claim 6, wherein determining
2 the conversion function comprises determining the

3 function so as to generate a data element indicated by
4 the corresponding one of the markup tags.

1 8. A method according to claim 6, wherein determining
2 the conversion function comprises determining the
3 function to generate an attribute of the unified data
4 indicated by the corresponding one of the markup tags.

1 9. A method according to claim 1, wherein at least some
2 of the source data are represented in a language other
3 than the markup language, and wherein mapping the source
4 data comprises transforming the data to the markup
5 language.

1 10. A method according to claim 1, and comprising
2 querying the sources by addressing a query to the unified
3 data in the markup language.

1 11. A method according to claim 10, wherein mapping the
2 source data comprises mapping the source data responsive
3 to the query.

1 12. Apparatus for processing source data from a
2 plurality of diverse sources in a selected data domain,
3 comprising a data integration processor, which is adapted
4 to receive and store a unified schema that lists markup
5 tags in the selected data domain that can exist in a
6 document in the markup language, and further to receive
7 and store definitions of correspondences of data fields
8 from the sources to the markup tags listed by the schema,
9 and to map the source data in accordance with the
10 correspondences to generate unified data in the markup
11 language.

1 13. Apparatus according to claim 12, wherein the markup
2 language comprises Extensible Markup Language (XML).

1 14. Apparatus according to claim 13, wherein the unified
2 schema comprises a Document Type Definition (DTD).

1 16. Apparatus according to claim 15, wherein the
2 processor is adapted to map the source data by
3 transforming the data using an XSL engine.

1 18. Apparatus according to claim 12, wherein at least
2 some of the source data are represented in a language
3 other than the markup language, and wherein the processor
4 is adapted to transform the data to the markup language.

1 20. Apparatus according to claim 19, wherein the
2 processor is adapted to map the source data responsive to
3 the query.

[illegible]